

A Critique of *Recognizing Digitization as a Preservation Reformatting Method*

by Andrew Hart

Andrew Hart (ashart@email.unc.edu) is Preservation Librarian, University of North Carolina at Chapel Hill.

After reading successive drafts and the final version of *Recognizing Digitization as a Preservation Reformatting Method*, I am troubled by a dissonance between the title and the actual messages of this document. Much more than “recognizing,” the Association of Research Libraries (ARL) Preservation of Library Materials Committee “endorses digitization as an accepted preservation reformatting option for a range of materials.” Yet, the Committee does not adequately address concerns about preservation implications inherent to digital reformatting – concerns that have been a frequent topic of discussion within

the profession over the past decade. This is not to say that digitization has no role in preservation but rather to say its role is too easily misunderstood. I believe ARL would better serve its membership and the field in general by exploring more systematically the conditions under which digitization is a plausible preservation strategy, describing the ongoing costs and risks in greater detail, and articulating more clearly the relationships between digitization for the purpose of preservation and digitization for other reasons.

Within librarianship’s overall goals of providing information resources to patrons who want

and need them, the defining aspect of preservation is attention to longevity of access. For any given information object, a number of strategies might be employed for this purpose. Whether by treating the artifact directly, providing surrogates to reduce wear on the original, replacing an original with a durable facsimile, managing storage and use environments, pursuing education and outreach, or planning for disaster response, the common thread in all preservation activity is the goal of reducing the risk of losing access to information. As an integral part of overall library goals of access and service, preserva-

tion is most successful when risk abatement actions are complementary with actions that improve patron satisfaction or that enhance use. Digitization holds the attraction of vastly improving patron satisfaction. It is problematic, however, because it carries a higher risk of loss than other approaches to preservation.

The ARL statement's assertion that "the time is right to adopt digitization as a reformatting strategy for preservation" is predicated on several assumptions: that the risk of catastrophic loss is acceptably low; the urgency of establishing sustainable long-term digital stewardship practices will successfully drive establishment and adoption of standards; and librarians will be most effective in meeting preservation goals by participating actively in standards development and learning from practice. Each of these assumptions is debatable and, taken together, they represent a radical change in how our profession thinks about and pursues the work of preservation.

Traditionally, much of what we do in preservation springs from how the physical vessels for information deteriorate. When something is weak, we strengthen it. When we observe harmful chemical reactions, we counteract them by controlling catalysts (e.g., heat and moisture) or by applying chemical treatments (e.g., deacidification). When an object is not likely to last long enough in one format, we do our best to transfer information to another, more durable, format. Such reformatting is expected to produce an access mode that requires no further

intervention to serve longer than the usable life span of the original.

Digitization is a reformatting method and, to that extent, builds on the precedents of preservation microfilming and photocopying and a variety of processes for reproducing non-print content. However, while the act of migrating content to fresh media is familiar, there are ways in which digitization represents a major shift in expectations for preservation. Rather than provide access media that are durable and comparatively independent of ancillary systems, digitization yields media that are known to be physically unstable and highly dependent on specific technological infrastructure. For example, we do not expect magnetic media to remain readable as long as even the most acidic piece of paper. Even when electronic storage media are engineered for greater physical durability, rapid cycles of hardware and software obsolescence render information inaccessible unless it is continually migrated to new systems. These concepts have been explored at great length in professional literature and conferences and have shaped the initiatives listed in the ARL endorsement.

Returning to the underlying assumptions in the endorsement, the authors seem to suggest that the risk of catastrophic loss is acceptably low because many capable people with support from major funding and research organizations are working very hard to establish sustainable strategies for information life-cycle management. Yet, in even the best-case scenarios under

discussion, we have to take a much bigger leap of faith in future development than ever before. In the past, the success of a preservation treatment has relied on future generations for little beyond keeping the treated material sheltered and not throwing it away. This passive baseline could be maintained whether or not anyone thought much about it. In contrast, for the foreseeable future we expect digital formats will require repeated and deliberate action from our successors to prevent irreversible loss. Furthermore, we expect digital formats will be very unforgiving of neglect during periods when some content has lower perceived value or when financial resources prevent adequate attention to migration.

The mainstream of librarians within ARL and elsewhere continues to take the position that the risk of losing born-digital content is a serious problem demanding our full attention because the universe of information in this category is growing quickly in quantity and complexity while we are still at a stage of developing strategies for long-term stewardship. Paradoxically, the Association is taking the position that the urgency of this problem makes it acceptable to exacerbate it. The reasoning seems to be that the risks described above are so serious that we will have to address them and therefore we can, in the name of preservation, add to the corpus of resources that must be preserved in digital form.

The idea that librarians will be most effective in meeting preservation goals by participat-

ing actively in standards development and learning from practice is compelling, but only up to a point. As with the precedents of microfilming and preservation photocopying, working experience will more than likely reveal problems and lead to improvements as any emerging preservation strategy matures. However, the lessons of microfilm and photocopies do not fully apply to digitization in some important ways, most notably in consequence of error.

Print on paper, microfilm, and xerographic reproduction have all been forgiving technologies. Even when we made poor choices of materials or failed to carry out a process according to today's benchmarks, we have had relatively ample opportunity to recognize and respond to our mistakes. "Slow fires" are, fortunately, slow. Just as important as the relative rates of deterioration, preserving the content of a brittle book or even a poorly processed reel of microfilm does not require that we preserve very much, if any, technological infrastructure from these objects' creation. In contrast, a digital file is not only subject to physical degradation of its storage medium but also to changes in the availability of hardware, software, and users' knowledge that can render it inaccessible. Not simply harder to read but completely useless.

The ARL endorsement makes questionable use of two quotations to support the contention that "the time is right to adopt digitization as a reformatting strategy for preservation." The first is an excerpt from an email from Abby Smith of the Council on Library and Information Re-

sources (CLIR). Smith writes, "As more and more is born digital and a new generation of users grows up with digital as the default mode of delivery, resources that are not in digital form will be 'orphaned' over time because they are in 'obsolete' formats." Smith's use of quotation marks suggests that she is not using the word "obsolete" literally in her reference to library users' strong and growing preference for digital delivery modes, often to the extent of ignoring valuable resources that are in less convenient formats. In this case, obsolescence is a reflection of information-seeking behavior rather than a change in the availability and utility of older media. This is an important distinction because data storage and retrieval systems have the potential to become technologically obsolete in ways that render large bodies of information completely inaccessible, which is clearly an even greater "inconvenience" than access in a less desirable format. While Smith's comment is a compelling argument in favor broadening the use of digital delivery modes for library content, it does not address the weaknesses of digitization as a *preservation strategy*.

The narrative portion of the ARL statement ends with another quote:

Libraries are society's stewards of cultural and intellectual resources. For libraries to continue fulfilling their stewardship role, they will have to approach preservation in a new way. It must be integrated into every aspect of the library's work. Preservation must be considered at the highest levels of the institution and reconceived in the digital environment. (From

the Preface to *The State of Preservation Programs in American College and Research Libraries: Building a Common Understanding and Action Agenda*, Deanna Marcum, President, Council on Library and Information Resources, December, 2002.)

Deanna Marcum wrote this passage for the preface to a survey report funded by the Institute of Museum and Library Services (IMLS) and co-sponsored by ARL, among other organizations. Based on my own experience conducting interviews and serving as an advisory committee member for this project, I believe relatively few libraries currently have the resident expertise, technological infrastructure, and financial resources necessary to employ digitization as a preservation reformatting method. In the CLIR study survey, most libraries reported having little or no plans for long-term digital stewardship and the qualitative segment of the project showed a thirst for standards, guidelines for best practice, and training. The context for Marcum's preface is the fourth recommendation in the CLIR report, which begins:

Of all the preservation challenges, none is more pressing than developing solutions to digital preservation. Staff members in academic libraries understand the general problem, but most do not know how to address it. (*The State of Preservation Programs in American College and Research Libraries: Building a Common Understanding and Action Agenda*, Anne R. Kenney and Deirdre C. Stam, Council on Library and Information Resources, December, 2002, page 9.)

All too often, libraries are struggling to provide adequate funding and leadership for core elements of their preservation

programs such as commercial binding, basic book repair and disaster response. The CLIR study cited in the ARL statement provides evidence that while preservation genuinely must be "reconceived in the digital environment," at present few libraries are ready to make a credible claim to using digitization for preservation purposes. With this in mind, I am concerned that ARL's endorsement may provide some libraries a rationalization for funding very attractive digital access projects in lieu of, rather than as an adjunct or integral part of, well-rounded preservation programs.

Without a doubt, digitization for a variety of purposes is a growing activity in libraries, already yielding tremendous benefits to library users. Recogniz-

ing the magnitude of unresolved questions as well as the promise for preservation, I propose a more moderate stance than ARL has presented, emphasizing the following priorities:

1. Establishing effective and widely-recognized technologies and business models for preserving digital content, including the products of reformatting projects as well as information that exists only in electronic formats;
2. Advocating for sufficient conservation of original artifacts as an aspect of digital reformatting projects;
3. Using currently available digital reformatting methods when they truly offer the best or only chance of survival for endangered information resources;
4. Promoting the use of standards and practices that are likely to improve the longevity of

digital objects created by reformatting.

We can serve our patrons best by making sure we understand clearly our reasons for pursuing digital reformatting, the risks and benefits in doing so, and the resources we will need for open-ended stewardship. The greatest gift of leadership by those at the cutting edge of practice is a message of caution. There is no question that digitization holds great promise for the future and many would agree that digital reformatting is already a viable preservation approach in some circumstances. However, it is both important and challenging to describe such circumstances and define the criteria that must be met in order to construe digitization as preservation.